

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2002, 03:17:36 ; Search time 39 Seconds
(without alignments)
202.064 Million cell updates/sec

Title: US-09-895-298a-83

Perfect score: 1002

Sequence: 1 MMNFQPPSKAWRASQMTFF.....HDSLDLRSSRVQEGNPRA 190

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|------------|--------------------|
| 1 | 99.5 | 9.9 | 1203 | YT41_CAEEL | Q11069 caenorhabd1 |
| 2 | 84 | 8.4 | 399 | YXAM_BACSU | P42112 bacillus su |
| 3 | 79.5 | 7.9 | 149 | Y38A_MYCGE | Q92b71 mycoplasma |
| 4 | 79 | 7.9 | 239 | PNUC_SALTY | P24520 salmonella |
| 5 | 77 | 7.7 | 342 | SPAS_SHIFL | P40707 shigella fl |
| 6 | 77 | 7.7 | 564 | YEDQ_ECO57 | Q8xb92 escherichia |
| 7 | 76.5 | 7.6 | 220 | T2N5_NOSS7 | P35677 nostoc sp. |
| 8 | 76 | 7.6 | 564 | YEDQ_ECOLI | P76330 escherichia |
| 9 | 75.5 | 7.5 | 1189 | ALAB_ARATH | Q91k90 arabidopsis |
| 10 | 75.5 | 7.5 | 2410 | MOKI_SCHPO | Q9usk8 schizosach |
| 11 | 75 | 7.5 | 422 | PANL_HUMAN | Q96rd7 homo sapien |
| 12 | 75 | 7.5 | 784 | TLR2_BOVIN | Q951a9 bos taurus |
| 13 | 75 | 7.5 | 784 | TLR2_HUMAN | Q60603 homo sapien |
| 14 | 74.5 | 7.4 | 148 | DSBI_BACSU | O64037 bacillus su |
| 15 | 74.5 | 7.4 | 385 | Y464_MYCGE | P47702 mycoplasma |
| 16 | 74.5 | 7.4 | 385 | Y464_MYCPN | P75112 mycoplasma |
| 17 | 74.5 | 7.4 | 401 | OPRM_BOVIN | P79350 bos taurus |
| 18 | 74.5 | 7.4 | 433 | YBL4_YEAST | P38211 saccharomyc |
| 19 | 73.5 | 7.3 | 159 | LSPA_STACA | Q59835 staphylococ |
| 20 | 73.5 | 7.3 | 616 | YGO2_YEAST | P30777 saccharomyc |
| 21 | 73 | 7.3 | 284 | COX3_LEITA | P14546 leishmania |
| 22 | 73 | 7.3 | 466 | GAB3_DROME | Q08832 drosophila |
| 23 | 73 | 7.3 | 516 | CALR_RAT | P32214 rattus norv |
| 24 | 72.5 | 7.2 | 568 | PTLB_STRMU | P50976 streptococc |
| 25 | 72.5 | 7.2 | 830 | YJ2_YEAST | P40367 saccharomyc |
| 26 | 72 | 7.2 | 362 | AOX_NEUCR | Q01355 neurospora |
| 27 | 72 | 7.2 | 444 | CYB_RHOSH | Q02761 rhodobacter |
| 28 | 71.5 | 7.1 | 500 | CLSL_BACSU | P45860 bacillus su |
| 29 | 71.5 | 7.1 | 971 | Y277_MYCPN | P75387 mycoplasma |
| 30 | 71.5 | 7.1 | 1064 | CY44_RAT | P26770 rattus norv |
| 31 | 71.5 | 7.1 | 1087 | AKA9_RABIT | Q28628 oryctolagus |
| 32 | 71 | 7.1 | 189 | Y78B_METJA | P81231 methanococc |
| 33 | 71 | 7.1 | 239 | PNUC_ECOLI | P31215 escherichia |

| | | | | | | |
|----|------|-----|------|---|------------|--------------------|
| 34 | 70.5 | 7.0 | 208 | 1 | ATPF_MYCGE | P47643 mycoplasma |
| 35 | 70.5 | 7.0 | 390 | 1 | EMRA_HAEIN | P44928 haemophilus |
| 36 | 70.5 | 7.0 | 398 | 1 | OPRM_MOUSE | P42866 mus musculu |
| 37 | 70.5 | 7.0 | 398 | 1 | OPRM_RAT | P33535 rattus norv |
| 38 | 70.5 | 7.0 | 400 | 1 | OPRM_HUMAN | P35372 homo sapien |
| 39 | 70.5 | 7.0 | 689 | 1 | DNLJ_RICPR | Q92k15 heliobacte |
| 40 | 70 | 7.0 | 226 | 1 | ATP6_HELPJ | Q56085 heliobacte |
| 41 | 70 | 7.0 | 226 | 1 | ATP6_HELPJ | Q10907 caenorhabd1 |
| 42 | 70 | 7.0 | 397 | 1 | YWO4_CAEEL | P44927 haemophilus |
| 43 | 70 | 7.0 | 510 | 1 | EMRB_HAEIN | P76007 escherichia |
| 44 | 70 | 7.0 | 578 | 1 | YCGO_ECOLI | P21860 homo sapien |
| 45 | 70 | 7.0 | 1342 | 1 | ERB3_HUMAN | |

ALIGNMENTS

| | | | | | | | | | |
|--|--|---|------|------|-----|--|--|--|--|
| RESULT 1 | | | | | | | | | |
| ID | YT41_CAEEL | STANDARD; | PRT; | 1203 | AA. | | | | |
| AC | Q11069; | | | | | | | | |
| DF | 01-NOV-1997 (Rel. 35, Created) | | | | | | | | |
| DT | 16-OCT-2001 (Rel. 40, Last sequence update) | | | | | | | | |
| DT | 16-OCT-2001 (Rel. 40, Last annotation update) | | | | | | | | |
| DE | Hypothetical 136.1 kDa protein B0416.1 in chromosome X. | | | | | | | | |
| GN | B0416.1. | | | | | | | | |
| OS | Caenorhabditis elegans. | | | | | | | | |
| OC | Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; | | | | | | | | |
| OC | Rhabditidae; Peloderinae; Caenorhabditis. | | | | | | | | |
| OX | NCBI_TaxID=6239; | | | | | | | | |
| RN | [1] | | | | | | | | |
| RP | SEQUENCE FROM N.A. | | | | | | | | |
| RC | STRAIN=Bristol N2; | | | | | | | | |
| RA | Favali T.; | | | | | | | | |
| RL | Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases. | | | | | | | | |
| RN | [2] | | | | | | | | |
| RP | REVISIONS. | | | | | | | | |
| RA | Waterston R.; | | | | | | | | |
| RL | Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases. | | | | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | | | | | |
| CC | ----- | | | | | | | | |
| DR | EMBL: U23516; AAC38884.1; . | | | | | | | | |
| DR | WormPep; B0416.1; CE19663. | | | | | | | | |
| KW | Hypothetical protein. | | | | | | | | |
| SQ | SEQUENCE 1203 AA; 136059 MW; 1B96886F827CC206 CRC64; | | | | | | | | |
| Query Match 9.9%; Score 99.5; DB 1; Length 1203; | | | | | | | | | |
| Best Local Similarity 27.7%; Pred. No. 0.26; | | | | | | | | | |
| Matches 39; Conservative 23; Mismatches 50; Indels 29; Gaps 8; | | | | | | | | | |
| QY | 6 | PPSKAWRASQMTFFIFLFFPSFTGLCTLAI--TIWRLKPSADCGPRLPLFIHSIY | 63 | | | | | | |
| DB | 701 | PASQIFRASRSNFFALLILFLF--LCFLPVGVFIASKTPSKSCGPF-GNQSFFYSVI | 756 | | | | | | |
| QY | 64 | SWIDTLSTRPGYLMVWVIYRNLSVHFF-----ILTLVIITTYLWQIT-----EG | 112 | | | | | | |
| DB | 757 | T--DYVHEN-----LDKTLVNGIKYLSPIIIVLVLSVITYFLIAMVTGLSQAN | 806 | | | | | | |
| QY | 113 | RKIMIRLLHEQIINEGRDKMF | 133 | | | | | | |
| DB | 807 | QDLSPQLMVER--TEEKKKIF | 825 | | | | | | |
| RESULT 2 | | | | | | | | | |
| YXAM_BACSU | YXAM_BACSU | STANDARD; | PRT; | 399 | AA. | | | | |
| ID | YXAM_BACSU | STANDARD; | PRT; | 399 | AA. | | | | |


```

FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 69 89 POTENTIAL.
SQ SEQUENCE 149 AA; 17731 MW; F0A7ABDDDD2562384 CRC64;

Query Match 7.9%; Score 79.5; DB 1; Length 149;
Best Local Similarity 24.7%; Pred. No. 1.7;
Matches 36; Conservative 22; Mismatches 53; Indels 35; Gaps 7;

OY 11 WRASQMTEFFIFLLFPP-SFTGLCTLAITWRKPSADCGPFRGLPIFHSISWIDTL 69
   | : : : | | : : | | : | | | | : | : | : |
DB 15 WLEFVPLLIYFLFAFLFAFTKYI-----IWELIPPCYSTAFTLIVFLSGIIP----- 63

OY 70 STRPGYLWVVIYRNILGSHVFFFIITLIVLIITL-----YQITEGRKIMIRLHE 122
   | | | : | | : | | | | : : : | : | | : |
DB 64 -----MANNSWI-----ILLRFVLVITLMLSLFLLNKMTNFFLIRSKYAITA---E 108

OY 123 QLINEGKDKMFLIEKLKLO-DMEKK 147
   : : | | : : | | : : | | : : |
DB 109 NLLKTGSKTKNRKLOQIOSDLERK 134

RESULT 4
PNUC_SALTY
ID PNUC_SALTY STANDARD; PRT; 239 AA.
AC P24520;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein pnuC.
GN PNUC OR STM0757.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RX [1]
RX RP SEQUENCE FROM N.A.
RX RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RA RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RA RT LT2.";
RA RN Nature 413:852-856(2001).
RR -1- FUNCTION: REQUIRED FOR NMN TRANSPORT ACROSS THE CYTOPLASMIC
RR CC MEMBRANE.
RR CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
RR CC -1- INDUCTION: REPRRESSED BY NADR.
RR CC -1- CAUTION: Ref.1 sequence differs from that shown due to a
RR CC frameshift in position 229.
RR CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
DR EMBL: M85180; -, NOT_ANNOTATED_CDS.
DR EMBL: AE008731; AAL19696.1; -.
DR PIR: D37753; D37753.
DR StyGene; SG10306; pnuC.

```

[illegible]

```

DR PIR: A38908; A38908.
DR InterPro; IPR002066; Bac_export_2.
DR Pfam; PF01312; Bac_export_2; 1.
DR PRINTS; PRO0950; TYPE3IMSPROT.
KW Virulence; Transmembrane; Inner membrane; Plasmid.
FT TRANSMEM 28 48 POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 181 201 POTENTIAL.
FT TRANSMEM 260 280 POTENTIAL.
SQ SEQUENCE 342 AA; 39852 MW; 48BBF531BFE034BC CRC64;

Query Match
Best Local Similarity 7.7%; Score 77; DB 1; Length 342;
Matches 44; Conservative 31; Mismatches 43; Indels 74; Gaps 11.

OY 22 FLLEFPSTFGVLCITAITWRLKPSA-----DCGPFKGLPLFHISYISWDTLSTPPGY 75
    1 1 1 1 : 1 1 1 1 : 1 1 1 1 1 1 1 1 1 1
DB 83 FPLEFCVLSAVLPTLVQTKFVLATKAKIDFSLNPKVL-----KKIFS----- 127
OY 76 LWWVWYIRNLIGSVHFF--ILTLIVL-IITYLXWQITEGRKIMIRLLHQIINEG---- 128
DB 128 -----IKTIKEFFKSILLIILALTYTFEW--INDRKIIFSQVFSSV--DGLYLI 173
OY 129 -----KDKM-----FLIEKLIKLODMKKANPSSLYLERREYQOGFLHLG 169
DB 174 WGRLEFKDILLFLAFLSILVILDEVIEFILYMKDM-----MMDKQETIKREYI---- 220
OY 170 EHDGSLDLKSR 181
DB 221 EDEGHFETKSR 232

RESULT 6
YEDO_ECO57 STANDARD; PRT; 564 AA.
ID YEDO_ECO57
AC Q8XB92;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein yedQ.
GN YEDO OR Z3047 OR ECS2694.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
OC Escherichia.
OX NCBI_TaxID=83334;
OX 111
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RL 121
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Ikeda T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kubara S., Shiba T., Hattori M., Shingawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE YAIC / YFIN (E.COLI), YHCK (B.SUBTILIS)
CC FAMILY.

```

```
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      DR      EMBL; AE005417; AAC56970.1; -.
CC      DR      EMBL; AP002559; BAB36117.1; -.
CC      DR      InterPro: IPR000160; GGDEF.
CC      DR      Pfam; PF00990; GGDEF; 1.
CC      DR      SMART; SM00267; DUF1; 1.
CC      DR      TIGRfam; TIGR00254; GGDEF; 1.
CC      KW      Hypothetical protein; Transmembrane; Complete proteome.
CC      FT      TRANSMEM     20      40      POTENTIAL.
CC      FT      TRANSMEM     360     380      POTENTIAL.
CC      SQ      SEQUENCE     564 AA; 64287 MW; BEBC2286ADBAECB0 CRC64;
CC
CC      Query Match                                7.7%; Score 77; DB 1; Length 564;
CC      Best Local Similarity    26.0%; Pred. No. 13;
CC      Matches    38; Conservative    25; Mismatches    43; Indels    40; Gaps    8;
CC
OY      64  SWIDTLSTR--PGYLVWVWIYRNLSGVHFFPILTYLYITYWQ---ITEGRKIMIR 118
DB      11  SWLKRLARLGPGR-----VNLCFI--VVLLFSTLTWRREVVLLEDAVISSQ 56
OY      119 LHEHQIINEGDKD-MELIEKLKLODMEKKA----NPSSLVERREVEQGQFLHGEH 171
DB      57  RNHLLENVANALDKHLQYNVDKLIIFLRNGMRREALVPAPLDFTSLRDVAITEREQ---HRDEH 112
OY      172 DGSIDLRSRRS-----VOEGN 187
DB      113 AMQIELNRRTLSVNGVS DALVSEGN 138
CC
RESULT 7
T2N5_NOSS7
ID      T2N5_NOSS7          STANDARD;          PRT;       220 AA.
AC      P35677;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      Type II restriction enzyme Nspv (EC 3.1.21.4) (Endonuclease Nspv)
DE      (R.Nspv).
GN      NSPVR.
OS      Nostoc sp. (strain PCC 7524).
OC      Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX      NCBI_TaxID=28072;
RX      [1]
RX      MEDLINE=93376523; PubMed=8367311;
RA      Ueno T., Ito H., Kotani H., Nakajima K.;
RT      "Cloning and expression of the Nspv restriction-modification genes of
RT      Nostoc sp. strain PCC7524.";
RL      Nucleic Acids Res. 21:3899-3899(1993).
CC      -1- FUNCTION: RECOGNIZES THE DOUBLE-STRANDED SEQUENCE TTGCAG
CC      AND CLEAVES AFTER T-2.
CC      -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage of DNA to give
CC      specific double-stranded fragments with terminal 5'-phosphates.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D14719; BAA03538.1; -.
DR      REBASE; 1407; Nspv.
KW      Hydrolase; Endonuclease; Nuclease; Restriction system.
```



```
FT DOMAIN 121 299 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 300 321 POTENTIAL.
FT DOMAIN 322 358 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 329 359 POTENTIAL.
FT DOMAIN 377 920 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 921 940 POTENTIAL.
FT DOMAIN 941 954 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 955 974 POTENTIAL.
FT DOMAIN 975 1004 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1005 1027 POTENTIAL.
FT DOMAIN 1028 1040 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1041 1063 POTENTIAL.
FT TRANSMEM 1064 1069 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1070 1090 POTENTIAL.
FT DOMAIN 1091 1107 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1108 1132 POTENTIAL.
FT DOMAIN 1133 1189 CYTOPLASMIC (POTENTIAL).
FT MOD_RES 424 424 PHOSPHORYLATION (BY SIMILARITY).
FT METAL 865 865 MAGNESIUM (BY SIMILARITY).
FT METAL 869 869 MAGNESIUM (BY SIMILARITY).
SQ SEQUENCE 1189 AA; 135309 MW; DCF951CF3FC55E83 CRC64;
```

Query Match 7.5%; Score 75.5; DB 1; Length 1189;
Best Local Similarity 23.4%; Pred. No. 40;
Matches 44; Conservative 30; Mismatches 71; Indels 43; Gaps 9;

```
OY 6 PSKAWRASQMTFFIFLFPSTFGLCTLAIT-----TWRLKPSADCGPFRGLPLFI 59
DB 285 PSKRSKIEKKMDQIITLFI-----SLIVIAFTGVSFEGIATRRDMSDNKRLR----- 333
OY 60 HSIYSWIDTSTRGYLMVWYIYRNLSGVHFFILITIV---LITTYLWQITEGRKIM 116
DB 334 ---RWY---LRDHTTFEYDPRRAVAAAFHFPLTALMLXGLIPISLYSIE-----V 380
OY 117 IRLHEQIINEGKDKMFLIKLQDMEKKNPSSLYERREVEQGGFLHGEHDSLD 176
DB 381 VKVLQSTIFINQDE-MYHEE-----TDRPARARTSNLNEELGOVD---TILSDKTGTLT 430
OY 177 LRSRSYQ 184
DB 431 CNSMEFVK 438
```

RESULT 10
MOK1_SCHPO STANDARD; PRT; 2410 AA.
AC Q9USK8; Q9UR75; Q9URK2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cell wall alpha-1,3-glucan synthase moki (EC 2.4.1.183).
GN MOK1 OR AGS1 OR SPC338.01C OR SPC17A7.01.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RA Katayama S., Dai H., Arellano M., Perez P., Toda T.,
RT "Fission yeast alpha-glucan synthase Moki localizes closely with actin
RT and play a role essential for cell morphogenesis and protein kinase C
RT function."
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgourou J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltham T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,

RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
RA Coffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottler S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gallard C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
RT "The genome sequence of Schizosaccharomyces pombe."
RL Nature 415:871-880(2002).
CC -I- CATALYTIC ACTIVITY: UDP-glucose + {alpha-D-glucosyl-(1,3)}(N) -
CC UDP + {alpha-D-glucosyl-(1,3)}(N+1).
CC -I- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AB019183; BAA34054.1; -
CC DR EMBL: AL023781; CAI19332.1; -
CC DR EMBL: AL049472; CAB39330.1; -
CC DR InterPro: IPR000461; Alpha_amyase.
CC DR InterPro: IPR001296; Glycos_transf_1.
CC DR Pfam: PF00128; alpha-amyase; 1.
CC DR Pfam: PF00534; Glycos_transf_1; 1.
CC KW Cell wall; Transferase; Glycosyltransferase.
CC FT CONFLICT 256 256 F -> Y (IN REF. 2).
SQ SEQUENCE 2410 AA; 272120 MW; 8BD944BAE3A9A5C5 CRC64;

Query Match 7.5%; Score 75.5; DB 1; Length 2410;
Best Local Similarity 24.3%; Pred. No. 87;
Matches 33; Conservative 13; Mismatches 49; Indels 41; Gaps 6;

```
OY 7 PSKAWRASQMTFFIFLFPSTFGLCTLAITWRLKPSADCGPFR 53
DB 2153 PSITWPLSAVSILIFALPKGLPEYRQLSGSIPAFYKSLRLRLVW----- 2200
OY 54 GLPLFHSIY---SWIDTSTRP-GYLWVWYIYRNLSGVHFFILITLYLYLTXWQI 109
DB 2201 ---FCTISVFLOQNWLSLNGRWSYLMW-----IGNIHQWQIFLLIYAFYIVL-WAL 2248
OY 110 TEGRKIMIRLHEQII 125
DB 2249 LLGLVLAWSIRTHSWII 2264
```

RESULT 11
PAN1_HUMAN STANDARD; PRT; 422 AA.
AC Q96RD7; O75968; Q96RS5; Q96L77; Q96AM9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pannekin 1.
GN PANX1 OR MRS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.


```
FT REPEAT 335 358 LRR 10.
FT REPEAT 359 384 LRR 11.
FT REPEAT 386 411 LRR 12.
FT REPEAT 412 436 LRR 13.
FT REPEAT 438 456 LRR 14.
FT REPEAT 457 476 LRR 15.
FT REPEAT 477 499 LRR 16.
FT REPEAT 501 521 LRR 17.
FT REPEAT 523 546 LRR 18.
FT DOMAIN 639 784 TIR.
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 784 AA; 90204 MW; 59FDB9DFA7984C18 CRC64;

Query Match 7.5%; Score 75; DB 1; Length 784;
Best Local Similarity 23.0%; Pred. No. 28;
Matches 45; Conservative 23; Mismatches 58; Indels 70; Gaps 11;

QY 21 IFLFFPSFTGVCTLATITW-----RLKPSADCGPFRGLPLFTHSIYS----- 64
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 598 LFLLL--LTGVLCRHFGHLMYMKMMWMLQAKRKPR--APRRDICYDAFVYSERDSY 653
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 65 WIDTLSTR-----PGYLVWVIYRNIGSV--HFFELTTLVLI 102
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 654 WVENLWVQLEHNPFPKLCILHKRDFIPG---KWIDNIIDSIEKSH---KTIFVLS 705
   ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 103 TYLYWQITGGRKIMIRLLHQIINECKDKMFLIKLQDMKKANPSSVLER--- 158
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 706 NFVK---SEWCKEELDFSHRFLFENDNDAT---LILLEPIDKKAIPQRCCKLRKIMNT 758
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 159 -----EVEOGEF 165
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 759 KTYLEWPEVDETQOGEF 774
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 13
TLR2_HUMAN STANDARD; PRT; 784 AA.
ID TLR2_HUMAN 015454;
AC 060603;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toll-like receptor 2 precursor (Toll/interleukin 1 receptor-like protein 4).
DE TLR2 OR TLR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte, and Prostate;
RX MEDLINE=98261424; PubMed=9596645;
RA Chaudhary P.M., Ferguson C., Nguyen V., Nguyen O., Massa H.F., Eby M.,
RA Jasmin A., Trask B.J., Hood L., Nelson P.S.;
RT "Cloning and characterization of two Toll/Interleukin-1 receptor-like
RT genes TLR3 and TLR4: evidence for a multi-gene receptor family in
RT humans.";
RL Blood 91:4020-4027(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98118556; PubMed=9435236;
RA Rock F.L., Hardiman G., Timans J.C., Kastelein R.A., Bazan J.F.;
RT "A family of human receptors structurally related to Drosophila
RT Toll.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:588-593(1998).
RN [3]
RP SEQUENCE FROM N.A., AND RESPONSE TO LIPOPOLYSACCHARIDE.
RC TISSUE=Fetal Lung;
RX MEDLINE=98421677; PubMed=9751057;
RA Yang R.-B., Mark M.R., Gray A., Huang A., Xie M.H., Zhang M.,
RA Goddard A., Wood W.I., Gurney A.L., Godowski P.J.;
```

```
RT "Toll-like receptor-2 mediates lipopolysaccharide-induced cellular
RT signalling.";
RT Nature 395:284-288(1998).
RN [4]
RP RESPONSE TO BACTERIAL LIPOPROTEINS.
RX MEDLINE=99357867; PubMed=10426996;
RA Aliprantis A.O., Yang R.-B., Mark M.R., Suggett S., Devaux B.,
RA Radolf J.D., Klimpel G.R., Godowski P.J., Zychlinsky A.;
RT "Cell activation and apoptosis by bacterial lipoproteins through
RT Toll-like receptor-2.";
RL Science 285:736-739(1999).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF TIR DOMAIN, AND MUTAGENESIS.
RX MEDLINE=20531768; PubMed=11081518;
RA Xu Y., Tao X., Shen B., Horing T., Medzhitov R., Manley J.L., Tong L.;
RT "Structural basis for signal transduction by the Toll/interleukin-1
RT receptor domains.";
RL Nature 408:111-115(2000).
CC -I- FUNCTION: Cooperates with MD-2 to mediate the innate immune
CC response to bacterial lipoproteins and other microbial cell wall
CC components. Acts via MyD88 and TRAF6, leading to NF-kappa-B
CC activation, cytokine secretion and the inflammatory response. May
CC also promote apoptosis in response to lipoproteins. Recognizes
CC mycoplasma macrophage-activating lipopeptide-2kd (MALP-2),
CC soluble tuberculosis factor (STF), phenol-soluble modulin (PSM)
CC and B.burgetteri outer surface protein A lipoprotein (Ospa-L)
CC cooperatively with TLR6.
CC -I- SUBUNIT: Binds MD-2 and TLR6 via the extracellular domain. Binds
CC MYD88 via their respective TIR domains.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -I- TISSUE SPECIFICITY: Highly expressed in peripheral blood
CC leukocytes, in particular in monocytes, in bone marrow, lymph node
CC and in spleen. Also detected in lung and in fetal liver. Levels
CC are low in other tissues.
CC -I- PTM: N-glycosylated (By similarity).
CC -I- SIMILARITY: BELONGS TO THE TOLL-LIKE RECEPTOR FAMILY.
CC -I- SIMILARITY: CONTAINS 1 TIR DOMAIN.
CC -I- SIMILARITY: CONTAINS 14 LEUCINE-RICH REPEATS (LRR).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF051152; AAC34377.1; -
DR EMBL; U88878; AAC34133.1; -
DR Genew; HGNC:11848; TLR2.
DR MIM; 603028; -
DR PDB; 1FYW; 22-NOV-00.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR000157; TIR_domain.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR Pfam; PF01582; TIR; 1.
DR SMART; SM00082; LRRCT; 1.
DR SMART; SM00369; LRR_TYP; 2.
DR SMART; SM00255; TIR; 1.
DR PROSITE; PS50104; TIR; 1.
KW Receptor; Immune response; Inflammatory response; Signal;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein;
KW 3D-structure.
FT SIGNAL 1 18 POTENTIAL.
FT CHAIN 19 784 TOLL-LIKE RECEPTOR 2.
FT DOMAIN 19 588 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 589 609 POTENTIAL.
FT DOMAIN 610 784 CYTOPLASMIC (POTENTIAL).
FT REPEAT 51 74 LRR 1.
FT REPEAT 75 98 LRR 2.
```



```
RESULT 15
Y464_MYCGE          STANDARD;          PRT;          385 AA.
ID   Y464_MYCGE
AC   P47702;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Hypothetical protein MG464.
GN   MG464.
OS   Mycoplasma genitalium.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=2097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 33530 / G-37;
RX   MEDLINE=96026346; PubMed=7569993;
RA   Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA   Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA   Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA   Nguyen D.T., Utterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA   Tomb J.-F., Dougherty B.A., Bort K.F., Hu P.-C., Lucier T.S.,
RA   Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RT   "The minimal gene complement of Mycoplasma genitalium.";
RL   Science 270:397-403(1995).
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC   -1- SIMILARITY: TO A M.CAPRICOLUM HOMOLOG.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC   or send an email to license@lsb-slb.ch).
CC   -----
DR   EMBL; U39728; AAC72484.1; -.
DR   TIGR; MG464; -.
DR   InterPro; IPR001708; 60kDa_innermeb.
DR   Pfam; PF02096; 60KD_IMP; 1.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM 39..59.. POTENTIAL.
FT   TRANSMEM 150..170.. POTENTIAL.
FT   TRANSMEM 222..242.. POTENTIAL.
FT   TRANSMEM 270..290.. POTENTIAL.
FT   TRANSMEM 325..345.. POTENTIAL.
SQ   SEQUENCE 385 AA; 44207 MW; BCB2964F7842893A CRC64;

Query Match          7 4%; Score 74.5; DB 1; Length 385;
Best Local Similarity 28.3%; Pred. No. 14;
Matches 41; Conservative 19; Mismatches 28; Indels 57; Gaps 9;

QY 56 PLFIHSIYSWIDTLSTPGYLMVW-----IY--RNLAGS--VHFFILTLVLI 101
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 103 PYTFSDY---TLAYGPFGVWPAQIVLPIMYATRPVLSGVELGFNMILSLIVL 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 102 ITLYWQITEGRKIMIRLLEHQQIINEGDKMFLEIKLIKLDMEKKANPSSLVLEKREVE 161
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 159 -----LLVR-----LITIVITLN-----STLALEKMN-E 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 162 QQGFLH--LGEHDSLDLRSRSRYQ 184
   | | | | : : : : | : : : | : : : |
DB 182 VQGIKLAELINAKYKGATDLSKRNRQ 206
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Search completed: November 9, 2002, 04:31:25
Job time : 42 secs